

# Dynamic Allostery: Linkers Are Not Merely Flexible

Buyong Ma,<sup>1</sup> Chung-Jung Tsai,<sup>1</sup> Türkan Haliloğlu,<sup>2</sup> and Ruth Nussinov<sup>1,3,\*</sup><sup>1</sup>Basic Science Program, SAIC-Frederick, Inc., Center for Cancer Research Nanobiology Program, NCI-Frederick, Frederick, MD 21702, USA<sup>2</sup>Polymer Research Center and Chemical Engineering Department, Bogazici University, Bebek-Istanbul 34342, Turkey<sup>3</sup>Sackler Institute of Molecular Medicine, Department of Human Genetics and Molecular Medicine, Sackler School of Medicine, Tel Aviv University, Tel Aviv 69978, Israel\*Correspondence: [ruthnu@helix.nih.gov](mailto:ruthnu@helix.nih.gov)

DOI 10.1016/j.str.2011.06.002

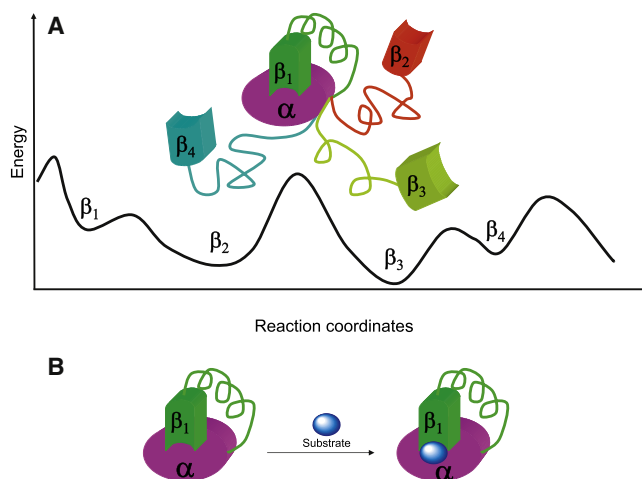
Most proteins consist of multiple domains. How do linkers efficiently transfer information between sites that are on different domains to activate the protein? Mere flexibility only implies that the conformations would be sampled. For fast timescales between triggering events and cellular response, which often involves large conformational change, flexibility on its own may not constitute a good solution. We posit that successive conformational states along major allosteric propagation pathways are pre-encoded in linker sequences where each state is encoded by the previous one. The barriers between these states that are hierarchically populated are lower, achieving faster timescales even for large conformational changes. We further propose that evolution has optimized the linker sequences and lengths for efficiency, which explains why mutations in linkers may affect protein function and review the literature in this light.

## Introduction

New proteins often evolve through expansion of the existing domain architectures, and functional complexity of proteins has largely been acquired through domain duplication and recombination (Cohen-Gihon et al., 2011). Domain composition and structural complexity correlate with biological processes, and domain rearrangements have had a key role in the emergence of typical features of vertebrates and chordates, in functional variation such as mating efficiency (Peisajovich et al., 2010), and in cellular pathway response. Protein domains are connected by linkers, indicating their importance (Wriggers et al., 2005). Because the majority of proteins have several domains, i.e., two-thirds of all proteins in prokaryotes and 80% in eukaryotes (Apic et al., 2003), considerable attention has focused on the properties of linkers and their roles. Early work demonstrated a relationship between linker flexibility and function (Gokhale and Khosla, 2000). Protein families were observed to undergo functionally relevant conformational changes that are similar (reviewed in Wriggers et al., 2005); and sequence analysis (George and Heringa, 2002) indicated that linkers vary in secondary structure and length (typically from ~5 to 25 amino acids) and often consist of flexible residues. Here, we argue that linkers are not merely flexible, and not only serve to prevent interdomain steric effects, because mere flexibility is unlikely to be sufficiently productive. From a functional standpoint the key point about linkers is in their allosteric role. The dynamics of linkers mediate the propagation force that originates from the perturbation caused by binding of ligands, or by covalent allosteric events such as posttranslational modifications occurring in one of the domains that they connect. The outcome is the fast reorientation of a second domain, e.g., a catalytic domain. Such a model implies that rather than just swivel and fluctuate, linkers encode a series of successive preferred states, in which each state encodes a subsequent one; i.e., although the functionally relevant orientations may represent rare high-energy states in the inactive protein, these states become more highly populated through allosteric propagation. The high flexibility of

the linkers implies that there are low barriers for the transitions between these states, thus, only short timescales between the allosteric event and its functional outcome. Figuring out these sequential preferred states that occur upon allosteric activation and force propagation is expected to help in understanding the conformational control of protein function and might also be useful in drug discovery. From a practical standpoint, this implies that allosteric drug discovery for multidomain proteins may benefit from targeting linkers (Liu and Nussinov, 2009, 2010a). Figure 1 presents an overview of such an allosteric view of the function of linkers.

Multidomain proteins are advantageous compared to associations of single proteins. This is because they increase the effective local concentration of substrates or products along enzyme metabolic or signaling pathways, which is expected to shorten the timescales of cellular response to environmental change. This may explain why during evolution, catalytic units that existed separately in simple organisms have been linked covalently (Marcotte et al., 1999). However, beyond the close physical confinement that avoids the time delay incurred by diffusion or collision of monomers (Echeverria and Kapral, 2010), or between reactant and products in subsequent enzymatic (Chen and Kapral, 2011) or signaling steps (Hollins et al., 2009), multidomain proteins allow to exert a more complex control. Proteins are regulated by transient interactions and covalent modifications. Allosteric propagation of the energy that is generated by such perturbation events via flexible linkers can lead not only to conformational changes of a second binding site in another domain but also to a relatively large, allosterically driven reorientation of protein domains with respect to each other (Liu and Nussinov, 2009; Zhuravleva and Gierasch, 2011). Here, we posit that efficient reorientation is not merely an outcome of global linker flexibility but that it relates to successive pre-encoded preferred dynamic states whose populations are governed by allosteric propagation (Figure 2). We further note that similar theoretical considerations of pre-encoded states apply to allosteric propagation pathways in any part of the structure;



**Figure 1. An Overview of the Linker Functions in Transmitting the Allosteric Propagation Force**

(A) In the unbound (inactive) state of the protein, the linker between domains  $\alpha$  and  $\beta$  fluctuates, and the  $\beta$  domain samples the conformational space, presenting certain populations of conformations:  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ , etc. Each conformation corresponds to an energy minimum on the energy landscape. The barriers separating the conformations can be high.

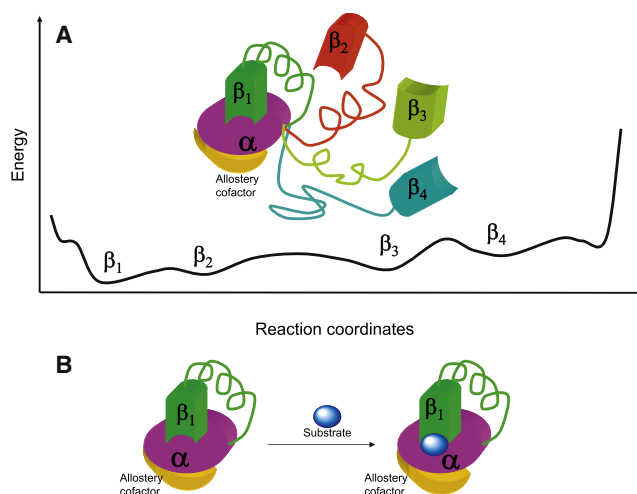
(B) For the protein to function, the  $\alpha$  and  $\beta$  domains must be in a certain orientation (conformation  $\beta_1$ ) with respect to each other, and the substrate-binding site must be in the “correct” conformation, which could be a high-energy state ( $\beta_1$ ). The linker is the string connecting the domains.

however, they may play a more pronounced role at those sites, which involve higher barriers and large conformational change.

### Modular Proteins Play Key Roles in Signaling Pathways via Their Linkers

Signaling pathways are crucial to the survival of cells and organisms (Wickström and Fässler, 2011), and allosteric propagation between protein domains is a hallmark of signaling pathways (Bruning et al., 2010; Brüscheweiler et al., 2009; Friedmann et al., 2008; Huang et al., 2011; Kannan et al., 2007; Kar et al., 2010; Petit et al., 2009; Zhang et al., 2010). Signaling involves a cascade of processes in which a stimulus is translated into cellular response. Most pathways are activated by external stimuli, and information is transferred from the cell surface to internal systems (Ma and Nussinov, 2009); however, some respond to information generated within the cell, usually in the form of metabolic messengers. Signaling may initiate with binding of relatively small molecules to extramembranous domains of receptors at the cell periphery with subsequent allosteric propagation through membrane-spanning domains accompanied by conformational changes (Tsai et al., 2009) through successive binding events in the cytoplasm to the nucleus; or by permeation of chemical messengers through the membrane. Cellular response can be reflected in changes in gene expression and degradation.

Signaling proteins are typically modular (Pawson, 2003). To transmit a signal along the pathway, the information needs to be communicated between the protein modules. By using modules, not only are catalytic activities regulated allosterically, but upstream and downstream partners can also be selected through allosteric conformational changes. Considering modules



**Figure 2. The Populations of Successive Pre-Encoded Preferred Dynamic States Can Be Governed by Allosteric Propagation**

(A) As the strain energy propagates, the linker undergoes a series of conformational changes, which redistribute the ensemble, and it progressively populates minor states, eventually leading to the functional, active state ( $\beta_1$ ). This redistribution populates states with lower barriers between them—not only compared to the preceding state but also to the following state. The sequence of the linker *pre-encodes* these series of states and as such has been engineered by evolution. As seen in the figure, following the series of states from  $\beta_1$  to  $\beta_{1,4}$  and the binding of the allosteric cofactor, barriers become progressively lower, leading to the active conformation ( $\beta_1$ ).

(B) In case of allosteric cofactor binding, the active conformation can be easily reached and bound by the substrate. Here, the substrate-binding site is between the two domains ( $\alpha$  and  $\beta$ ). However, in other cases the active conformation could be one, in which the two domains separated as shown in Figure 3.

as signaling units (Chang et al., 2009) raises the question of how the information is transferred from one module to the other. Because in single-chain proteins the modules are connected by linkers, linkers become the key; they are pivotal for efficient information transfer and for functional control, and they are responsible for coordination between the receiving (allosterically induced) module and the output module. Below, we argue that linkers encode *successive* conformational states, which predispose them to optimally fulfill their function. Thus, although all states preexist in the inactive protein, the relative distributions of the states change during an allosteric propagation in the linker.

### An Overview of the Free Energy Landscape, Allosteric Propagation, and Pre-Encoded States

In the cell, all (dynamic) proteins exist as ensembles of conformational states (or substates) around their native states that are in dynamic equilibrium (Frauenfelder et al., 1991). The distributions of the states are determined by statistical thermodynamics. The states are separated by barriers, the height of which defines the timescale for conformational exchanges. Proteins that are flexible have low barriers, which lead to a fast interconversion between the different states. For disordered proteins that have only small hydrophobic cores, the barriers are very low, and the population of each state is small and difficult to observe experimentally (Tsai et al., 2001). In all cases and for all proteins, all conformational states and substates preexist in the ensemble (Ma et al., 1999; Tsai et al., 1999b).

Allostery is regulation at a distance by conveying information from one site to another. The effector perturbs the structure of the first site and thereby leads to an altered activity in a second, substrate site (del Sol et al., 2009; Gunasekaran et al., 2004). Structural perturbations by changes in the environment, post-translational modification states, mutational events, or binding do not lead to new states; these events only change the relative distributions of the different states in the ensemble. The effect of the perturbation resembles disturbances of a water surface, which, if uneven, propagate as irregular waves that change in speed, frequency, and depth. The waves may merge and get stronger, or weaker if different frequencies and speed are involved. Allostery is similar to waves rippling outward from the points where the disturbances (altered atomic interactions) took place. Because proteins are tightly packed, atoms cannot move freely, but they can vibrate. Their collective energy travels, undulates, and oscillates in certain directions, and deforms the protein structure. The changes in the atomic interactions in the allosteric site caused by the binding or covalent modification lead to a local strain; and because the protein structure is nonhomogeneous, the allosteric strain energy generated at the perturbation site dissipates in the structure nonhomogeneously. Dissipation involves propagation of the changes in the atomic interactions to relieve the strain, and takes place through multiple, major and minor, pathways that depend on the protein topology and the distribution of the ensemble. Regardless of the location, stimulant, and type of the perturbation events (e.g., binding, mutations), no new pathways are created, but instead, the preexisting propagation pathways are always defaulted to. Allostery is a property of the ensemble and illustrates that the energy landscape is dynamic: perturbation at any site in the protein structure leads to a shift in the distribution of the preexisting conformational states (Cui and Karplus, 2008; del Sol et al., 2009; Formanek et al., 2006; Goodey and Benkovic, 2008; Gunasekaran et al., 2004). Such a thermodynamic description implies that allostery can be expressed in terms of changes in enthalpy, enthalpy and entropy, or solely entropy. That is, allostery may take place even in the absence of any observable conformational change, i.e., it can be either solely entropy driven (Tsai et al., 2008), or driven by entropy and enthalpy (Formanek et al., 2006; Itoh and Sasai, 2010; Okazaki and Takada, 2008; Tsai et al., 2008). Because all conformations preexist, binding events proceed via a conformational selection mechanism of the most geometrically and chemically favored conformers, even if they have higher energy (del Sol et al., 2009; Gunasekaran et al., 2004; Kenakin and Miller, 2010; Ma et al., 1999; Tsai et al., 1999a, 1999b; Zhuravlev and Papoian, 2010). Binding stabilizes these conformers, and the binding event is followed by a shift of the population toward these conformers. This phenomenon is the origin of the allosteric effect: the perturbation that occurs during the binding of the effector at the allosteric site propagates in the structure, leading to conformational (and/or dynamic) changes (Boehr et al., 2009; Kenakin, 2010; Keskin, 2007; Tobi and Bahar, 2005; Tsai et al., 2008). The changes can be minor, but they can also be amplified, resulting in the selection of different ligands. This mechanism, which is called “conformational selection and population shift,” was recently validated through an extensive compilation of experimental data (Boehr, 2009).

We proposed the concept of population shift of the ensemble following a perturbation event over a decade ago (Boehr et al., 2009; Ma et al., 1999; Tsai et al., 1999b). However, experimentally following specific pathways to demonstrate that it occurs on the atomic level has been challenging. Nuclear magnetic resonance (NMR) increasingly provides information on protein dynamics from which pathways can conceivably be inferred (Boehr et al., 2006, 2010; Kalodimos, 2011; Kern and Zuiderweg, 2003; Mittermaier and Kay, 2009; Swain and Gierasch, 2006; Volkman et al., 2001). Recently, using NMR dynamics and molecular dynamics simulations, an atomic scale pathway for a complex conformational transition was inferred (Boehr, 2009; Gardino et al., 2009). This allowed addressing the key question of how a population shift is achieved following perturbation. Allosteric propagation is the means by which a population shift takes place. Propagation implies crossing barriers, and because barriers imply high-energy states, it occurs through rare states being visited. As the allosteric strain energy is transferred along the pathway, barriers are crossed, which may facilitate experimental observation. In the particular case of NtrC<sup>r</sup>, conformational transition barriers were lowered through the formation of transient nonnative hydrogen bonds, which partly compensate for the loss of native contacts. However, other interactions may also be important for crossing energetic barriers, even if they are as yet unobserved (Boehr, 2009).

### Fluctuations and Timescales

At the one end of the spectrum, the interdomain motions can involve large-scale structural rotations and translations of domains at timescales of microsecond to millisecond; at the other end, there are small-scale structural transitions among an extremely large number of states that differ only in side-chain and small backbone dihedral angles that happen at fast timescales in the range of picosecond to nanosecond. These fast transitions normally display similar potential energies and very similar structures, and are separated by low barriers. Segmental and loop movements can be in-between. Larger interdomain rotations can be the outcome of linker flexibility and may involve high barriers. Here, we argue that the heights of these barriers are reduced as the allosteric propagation front, which initiates following an allosteric perturbation event, progresses. This dynamic progression shifts the populations in successive states favoring transition paths that circumvent high barriers toward functionally competent states, where the domains are “properly” oriented. It was proposed that the rate by which an enzyme converts substrates into products is not the speed of the chemical step (Villali and Kern, 2010). Instead, for efficiency, all steps need to be well executed because a single high barrier will slow down the entire reaction. Therefore, to avoid high barriers, catalytic efficiency is obtained through equilibrium fluctuations between consecutive, hierarchical states. Although more states may be occupied on the way to the catalytically optimal ones, the rate will be faster. On the other hand, it was argued (Whitford et al., 2008) that rather than obtaining the overall enzyme turnover rate from the conformational dynamics around the equilibrium basin (Villali and Kern, 2010), an allosteric change in protein structure can also control the process. According to this view, the energy surface of the protein is determined by the end structures resulting from the conformational change. A major issue is

whether the protein makes use of specific hinges or whether it “cracks and accesses partially unfolded states during its structural change” (Whitford et al., 2008).

Both descriptions could be general and hold true, not only for enzyme catalysis (Ma and Nussinov, 2010). Pre-encoded states have recently been observed by NMR and captured by molecular dynamics for dynamics-governed allostery of protein kinase A with minor conformational change (Masterson et al., 2011). Allosteric dynamic activation (Popovych et al., 2009; Tzeng and Kalodimos, 2009) is typically accompanied by minor conformational changes (Tsai et al., 2008).

### The Key Point: Linkers Encode Successive States

In the unbound (inactive) state of the protein, the linker fluctuates, and the domains sample the conformational space, presenting certain populations of conformations:  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , etc. (Figure 1A). However, for the protein to function, the domains need to be in a certain orientation ( $\beta_1$ ) with respect to each other (Figure 1B), which suggests that the linker must be in the functionally favored state. In the unbound state the population of this conformation can be low. Binding will cooperatively increase the population of this state; however, for a functionally swift response, the barriers leading to this state should be low. We suggest that the propagation of the strain energy from the perturbed binding site changes the relative populations of the states. The allosteric front will progressively redistribute the ensemble (Figure 2), which will be reflected in a series of observed conformational changes of the linker leading to the favorable  $\beta_1$  state (Zhu et al., 2011). The barriers between these states, which are hierarchically populated, are lower, achieving faster timescales even for large conformational changes. Different residue positions in the linker impact the linker conformation and dynamics in different ways, and local conformational changes within the linker can have profound effects on the transition pathway. The propagation energy can be concentrated at certain locations (Csermely et al., 2010; Piazza and Sanejouand, 2008, 2009), which can lead to large amplitude fluctuations of the protein domains (Haliloglu and Erman, 2009). Minor states can become more populated, and the threshold between these and subsequent functionally relevant states can be lower than between other states, which can allow response at a fast timescale. When we consider fast cellular response along a metabolic pathway that consists of many large, multidomain proteins, and of allosterically controlled catalysis, timescales are a key factor; and for short times, linkers should not fluctuate to populate mostly low-energy states. To circumvent a scenario in which the ensemble rarely visits linker functionally relevant near-native states, evolution appears to have encoded the linker sequence to successively follow the states that lead to the functional orientation. Allosteric propagation proceeds between states, which are separated by low thresholds. From the mechanistic standpoint, more favored states may be “locked” transiently by some stabilizing interactions, such as nestling of a linker into a hydrophobic pocket of one of the domains, or fleeting hydrogen bonds, which increases their populations (Figure 2). Because the transition pathway is important for protein function, residues at certain positions are conserved. Evolution makes use of the multiple states and substates for fast transmission to achieve the required conformational change.

Linkers can present large fluctuations of the protein domains, which may preclude structural determination of the entire protein, and necessitate to determine the structures of separate domains (George and Heringa, 2002). We suggest that in these cases there is a high cooperativity between the domains with a high probability of efficient signal response. pVHL and other substrate-binding proteins in the E3 ubiquitin ligase system can present good examples. Stabilizing such linkers may abolish biological function.

Although linker involvement often implies large conformational change, apparently this may not always be the case. In ligand-gated ion channels, the binding of a ligand to an intracellular or extracellular domain allosterically perturbs the transmembrane pore-forming helices, which leads to changes in ion flow. Fluorescent resonance energy transfer (FRET) between membrane-resident quenchers and fluorophores attached to the channel did not detect any movement orthogonal to the membrane during channel activation (Taraska and Zagotta, 2007). FRET measurements between fluorophores within the C-terminal region illustrated that when channels open, the C-terminal end of the C-linker and the end of the C helix move apart, suggesting that during channel activation, part of the gating ring moves in parallel with the membrane, toward the central axis of the channel. Overall, rather than a large conformational change, the movements within the C-linker appear to be subtle, involving only limited structural rearrangements (Taraska and Zagotta, 2007). The main observed movement was a separation of a loop between the ligand-binding domain (LBD) and the  $\alpha$  helices of the C-linker, from the end of the C helix. However, owing to the limited points of measurement, the large FRET distance (30–70 Å), and the relatively large probe sizes, these results may not be conclusive.

### Examples: Linkers Are the Vehicle for Transmission of Information Gene-Specific Transcription Factors

The glucocorticoid receptor (GR) can serve as one example for linkers that have allosteric roles. The GR consists of the N-terminal domain, DNA-binding domain (DBD), hinge region, LBD, and the C-terminal domain. Binding of agonists such as hormones to the LBD in the cytoplasm induces a major allosteric effect (Gronemeyer and Bourguet, 2009). The GR homodimerizes, translocating into the nucleus where finally the DBD binds to specific DNA response elements (REs) to activate transcription initiation. Selective binding is achieved through allosteric conformational changes of the DNA-binding surface of the DBD. In turn, DNA binding induces a conformational change of the lever arm, which alters the cofactor-binding sites, to modulate the glucocorticoid activity. Allosteric conformational changes can further alter GR surfaces that interact with coactivators including p160/SRC (steroid receptor coactivator) proteins, which can interact with histone acetyltransferases, such as CBP (cAMP response element-binding protein (CREB)-binding protein) and p300. The activated GR can also form complexes with other TFs and prevent their binding to target genes and, hence, repress the expression of genes that are normally upregulated by NF- $\kappa$ B or AP-1. The GR example illustrates how a receptor can fine-tune its target genes via



allosteric effects that are propagated through the linker. Of particular interest, even a single base pair change in the RE leads to altered functional effects (Meijsing et al., 2009) mediated by allosteric propagation. p53 provides another example of a signaling pathway in which allostery has a role. Binding of the p53 DBD to the p53 REs initiates signaling that propagates to the p53 activation domain (p53AD), which in turn binds the Mediator complex to activate or initiate transcription by RNA Polymerase II (Pol II) at the promoter (Meyer et al., 2010). Here, the DNA acts as an allosteric effector, and the signal propagates from the DBD to the activation domain through a flexible linker to result in binding to Mediator subunits (Guglielmi et al., 2004). Different REs undergo slightly different atomic contacts with the DBD, which result in different pathways that transmit DNA sequence specificity to the activation domain to activate an initiated-and-stalled Pol II (Meyer et al., 2010). A similar scenario takes place with other TFs, e.g., GCN4 (Natarajan et al., 1999; Swanson et al., 2003; Zhang et al., 2004), VP16 (Ito et al., 1999), C/EBP $\beta$  (Li et al., 2008), PGC-1 $\alpha$  (Wallberg et al., 2003), ER $\alpha$  and ER $\beta$  (Kang et al., 2002), RAR (Lee et al., 2007), HNF-4 (Malik et al., 2002), TR (Malik et al., 2004), Pit-1, GATA-1 and GATA-2 (Stumpf et al., 2006), PPAR $\gamma$  (Ge et al., 2008), GR (Chen and Roeder, 2007; Chen et al., 2006), and GABP (Udayakumar et al., 2006).

#### **A Light-Activated Prokaryotic Repressor**

When illuminated, 1 out of 12 rationally designed fusions between the naturally photoactive LOV2 domain from *Avena sativa* phototropin 1 and the *Escherichia coli* trp repressor selectively bound operator DNA and protected it from nuclease digestion, illustrating that an allosteric lever arm linker can couple the function of two domains. The designs sought a rigid  $\alpha$ -helical domain linker with defined geometry as an effective conduit for allosteric signals (Strickland et al., 2008).

#### **Neurotransmitter Receptors**

N-methyl-D-aspartate (NMDA) receptors (NMDARs) are excitatory neurotransmitters that form calcium-permeable, glutamate-gated ion channels that serve to mediate synaptic plasticity. There are multiple subtypes with distinct pharmacological and biophysical properties that are largely determined by the type of the NR2 subunit in the NR1/NR2 complex. A major difference between the subtypes is their channel maximal open probability, which spans a 50-fold range and confers unique charge transfer and signaling properties on each receptor subtype. Subunit-specific gating is controlled by the extracellular NR2 amino-terminal domain (NTD), which binds allosteric inhibitors and by the short linker through which the signals propagate to the remote agonist-binding domain (ABD). Subtype specificity largely reflects differences in the equilibrium between open- and closed-cleft conformations of the NR2-NTD (Gielen et al., 2009). The closed-cleft conformation triggers disruption of the ABD dimer interface and subsequent channel closure. In this allosteric, drug-based bidirectional control of the receptor activity, molecules can either promote inhibition by inducing closure of the NR2 channel by the NTD or potentiation, by helping to open it by a conformational change of the NR2/NTD region. The NR2-NTD region also determines the sensitivity to zinc and protons. Under physiological and pathological conditions, these act as endogenous allosteric inhibitors in the regulation of NMDAR activity.

#### **An Enzyme: Phosphodiesterase**

The structure of PDE includes catalytic and regulatory domains. PDE2A is a dimer in which each subunit had an extended organization of regulatory GAF-A and GAF-B and catalytic domains that are connected by long  $\alpha$  helices. The subunits cross at the GAF-B catalytic domain linker, with each half of the dimer containing the GAF-A and GAF-B domains of one subunit and the catalytic domain of the other subunit. In dimeric PDE2A the H-loops of the two catalytic subunits pack against each other at the dimer interface occluding the binding pocket, which necessitates the movement of the catalytic subunits to allow for H-loop movement. cGMP binding to GAF-B leads to a movement of the catalytic domains through the linker region, swinging the H-loops away from the dimer interface to allow substrate access. The increase in substrate access through the linker appears to be the basis for PDE2A activation by cGMP and a general mechanism for regulation of all PDEs (Pandit et al., 2009). The tandem GAF domain of hPDE10A uses cAMP as an allosteric ligand (Gross-Langenhoff et al., 2006). Changes at the C-6 ring position (removal of the amino group; chloride substitution) and at the N-1 ring position drastically reduced stimulation, marking those positions as being crucial to nucleotide discrimination. A GAF tandem chimera that consisted of the cGMP-binding GAF-A from hPDE5A1, which signals through cGMP in PDE5, and the GAF-B from hPDE10A1, which signals through cAMP in PDE10, illustrated that stimulation was through the hPDE10A1 GAF-B domain. Of particular interest, generation of additional PDE5/10 chimeras demonstrated that the length-conserved linker in GAF tandems between GAF-A and GAF-B played a key role in intramolecular signaling. Swapping linkers between PDE5 and PDE10 GAF tandem domains abolished signaling (Hofbauer et al., 2008).

#### **Allosteric Linkers in Molecular Machines**

Linkers are important for the efficient regulation of enzymes and play a key role in the regulation of complex molecular machines. The E3 ubiquitin ligase provides one example. The multimodule cullin-RING ligases (CRLs) resemble two-arm machines: the substrate binds to the substrate-binding protein, which connects to an adaptor protein on one arm, whereas the ubiquitin is covalently bound to the E2-conjugating enzyme that interacts with the Rbx protein on the other arm. The two arms are connected by the cullin, which has two domains. Piecing together all the available crystal structures has shown that the distance between the substrate-binding site and the E2 ubiquitin-binding site is 50–60 Å (Cardozo and Pagano, 2004; Schulman et al., 2000; Zheng et al., 2002). Experiment has shown that the linker is flexible (Duda et al., 2008), and simulations have illustrated that the linker between the substrate-binding domain and the box domain in the substrate-binding protein is allosterically regulated by the binding of either of the domains (Liu and Nussinov, 2009, 2010a), and similarly this is the case for the linker in the Rbx protein (Liu and Nussinov, 2010b) and likely also in cullin. The linkers not only fluctuate upon binding of either of the domains, or the covalent attachment of the Nedd8 allosteric effector to a conserved lysine in the C-terminal domain of the cullin, they demonstrate large rotations, which serve to provide a conformational control over the E3 ubiquitination machine, shrinking the distance and facilitating the ubiquitin-transfer reaction. Additional examples include the GRoEL, RNA Pol II, the hexameric

MecA-ClpC molecular machine with two nucleotide-binding domains (Wang et al., 2011), and likely most (if not all) other proteins that function as machines. Machines need to be regulated.

### Covalent Changes in Linkers: Phosphorylation, Mutations, and Alterations of Lengths

Covalent changes in the linkers shift the free energy landscape; as such, they generate allosteric effects similar to those induced by binding events to either domain (del Sol et al., 2009; Gunasekaran et al., 2004; Ma et al., 1999; Tsai et al., 1999a, 1999b). Covalent changes may include posttranslational modifications, mutations, and differences in linker length or sequence, as discussed below.

#### Posttranslational Modifications

Cbl protein ubiquitin ligases (E3s) play an important role in regulating tyrosine kinase signaling and provide an example for linker phosphorylation and allosteric consequences. The three mammalian family members, Cbl, Cbl-b, and Cbl-c, have a highly conserved N-terminal tyrosine kinase-binding domain, a catalytic RING finger domain, and a C-terminal proline-rich domain, which interacts with Src homology 3 (SH3)-containing proteins. The N termini of Cbl, Cbl-b, and Cbl-c inhibit their E3 activity, and the region responsible for this effect is the EF-hand and SH2 domains. However, phosphorylation of a key tyrosine (Tyr-341) in the linker region of Cbl-c (by Src) or a phosphomimetic mutation of this tyrosine (Y341E) can already increase the E3 activity of Cbl-c. These events also lead to a decrease in the affinity for the ubiquitin-conjugating enzyme (E2), UbcH5b, which leads to a more rapid turnover of bound UbcH5b (Ryan et al., 2010). Syk protein-tyrosine kinase that is phosphorylated on multiple sites after the aggregation of the B cell antigen receptor (BCR) provides an additional example. Among these, phosphorylation of serine-291 by protein kinase C enhances Syk's ability to couple the antigen receptor to the activation of the TFs NFAT and Elk-1. Serine-291 is in a 23 amino acid insert in the linker B region that separates the tandem pair of Src homology 2 domains from the C-terminal catalytic domain and distinguishes Syk from SykB and Zap-70. Removal of the entire linker insert did not block the transit of the kinase into the nucleus. However, nuclear import required a region near the C terminus of linker B, which suggested that it may not constitute a nuclear localization signal. Furthermore, whereas phosphorylation of Ser-291 did not affect the intracellular location of the kinase, replacement of Ser-291 by alanine decreased Syk's ability to couple the BCR to the NFAT (nuclear factor of activated T cells) activation, indicating decreased signaling. Overall, protein interaction studies have indicated that the phosphorylated linker insert helps in promoting an interaction between Syk and the chaperone protein prohibitin (Paris et al., 2010). Other posttranslational modification events such as glycosylation also take place in the linkers, as in the case of the *Trichoderma reesei* Family 7 cellobiohydrolase (Cel7A), in which linkers that are rich in serine and threonine with O-glycosylation connect the catalytic domains for cellulose hydrolysis and carbohydrate-binding modules. Molecular dynamics simulations focused on the resulting effects on the linker's hinge-bending motions, the relative orientation of the domains they connect, and their flexibility

or stiffness, and concluded that the linker is important in extending the distance between the domains and in this way expands the operating range of the enzyme (Beckham et al., 2010). Voltage-gated potassium Kv1 channels provide another example. These channels have three extracellular linkers that connect transmembrane domains, S1-S2, the S3-S4, and the S5-P, of which S1-S2 is the only one with a conserved N-glycan position in Kv1.1-Kv1.5 and Kv1.7 channels, shown to be essential for function through "structural restriction" (Zhu et al., 2009). Other posttranslational modification events, such as methylation, acetylation, and nitrosylation, can be expected to also play allosteric roles. Dynamic allosteric coupling of part of the S1 helix and the selectivity filter was also observed in a recent normal mode analysis of the Kv1.2 channel (Yehekel et al., 2010).

#### Mutations

Mutations can involve substitutions, deletion or insertion of residues, chimera, or linker deletion. The Tec kinase family is tyrosine kinases that function primarily in hematopoietic cells. The catalytic activity of their kinase domains is positively regulated by signals that are transmitted from the regulatory to the catalytic domains through a conserved spine. In the Tec kinases the regulatory spine, which assembles and disassembles as the kinase switches between its active and inactive states, includes a methionine in the C helix and a tryptophan in the Src homology 2-kinase linker. Mutation of the gatekeeper residue at the edge of the regulatory spine leads to a constitutively active kinase domain. The regulatory spine is preassembled in this gatekeeper mutant, rendering phosphorylation of the activation loop unnecessary for activity. Moreover, disruption of the conserved electrostatic interaction between Bruton's tyrosine kinase R544 on the activation loop and E445 on the C helix also helps with the assembly of the regulatory spine. Thus, the extended regulatory spine is a key structure that is critical for maintaining the activity of Tec kinases (Joseph et al., 2010). An example for mutations involving chimera is the chimeric protein LLhP, which comprises the LacI DBD, the LacI linker, and the PurR regulatory domain. Although DNA-binding site residues are identical in LLhP and LacI, LLhP does not discriminate between alternative DNA ligands as well as LacI. Two additional substitutions in LLhP diminished affinity, enhanced allostery, and altered DNA ligand selectivity, taken together, pointing to positions within the linker that can be varied to modulate repressor function (Zhan et al., 2008). The third example for the role of mutations relates to the effect of linker deletion. BzdR is a transcriptional regulator of the P(N) promoter that controls the anaerobic catabolism of benzoate in the *Azoarcus* sp. The effector molecule benzoyl-CoA was shown to induce conformational changes in BzdR without affecting its oligomeric state; however, the monomeric BzdR4 and the BzdR5 mutant showed that dimerization of BzdR is essential for DNA binding. Interestingly, the BzdRΔL protein, which lacks the linker region connecting the N- and C-terminal domains of BzdR, is also dimeric and is a super repressor of the P(N) promoter. Thus, the linker region of BzdR is not essential for protein dimerization; however, it is required for transmitting the conformational changes that are induced by the benzoyl-CoA to the DNA-binding domain, leading to the release of the repressor (Durante-Rodríguez et al., 2010).

### Difference in Linker Length

Comparison of domains of multidomain proteins that are involved in intracellular-signaling pathways revealed striking similarities across different species, which probably emerged prior to the last common ancestor of eukaryotes, bacteria, and archaea (Aravind and Subramanian, 1999). However, in contrast to the domain conservation, the lengths of the linkers between the domains in members of the family present large variability. This can often be seen in sequence alignments in which the sequences of the domains are fairly similar, but the linker regions manifest large insertions or deletions. A question that has largely been overlooked is why the linker length is so variable. One example is the smad ubiquitination regulatory factor 2 (Smurf2), which is an E3 ubiquitin ligase. Smurf2 plays a role in degradation of transforming growth factor (TGF)- $\beta$  receptors and other targets. Its WW domains recognize PPXY (PY) motifs present either on target proteins or on adapters such as Smad7. The isolated WW3 but not the WW2 domain of Smurf2 can bind directly to a Smad7 PY motif. However, WW2 strengthens this interaction by binding to the WW3 domain and making additional contacts with the PY and E/D-S/T-P motifs and enhances Smurf2 selectivity. Analogous WW domains in a short Smurf1 isoform recognize the Smad7 PY peptide using the same coupled mechanism. However, a longer Smurf1 isoform, with an additional 26 residues in the inter-WW domain linker, can only partially use this coupled WW domain-binding mechanism because the longer linker leads to a decrease in affinity for the Smad7 peptide (Chong et al., 2010). The length of linkers between pVHL ubiquitin ligase and estrogen receptor target protein significantly affects the efficiency of ubiquitination in a designed proteolysis that targets chimeras (Cyrus et al., 2011) through allosteric effect of binding to either of the linked proteins (Liu and Nussinov, 2010a). A cooperative allosteric effect was observed by linkers as short as two residues (Gly-Ala) in a crystalline porous solid design (Rabone et al., 2010). In another example, transmission of mechanical forces through the 14–18 amino acid long neck linkers in the kinesin-1 (Kin1) and kinesin-2 (Kin2) motors is the molecular basis that underlies their coordinated stepping. Of striking interest, the difference in processivity between the Kin1 and Kin2 motors was shown to result from differences in the length of the neck linker, and not from inherent differences in kinetic rates in the heads; that is, in Kin1 and Kin2 that have identical neck linker lengths, their run lengths match, even though their motor velocities differ by nearly a factor of two. Extension of the neck linker slows strain-dependent detachment of the rear head along with reduced strain-dependent inhibition of ATP binding to the leading head. Motors in kinesin families have different neck linker lengths, and they possess different processivities, which suggests that a correlation of neck linker length with function may extend across the kinesin superfamily (Shastri and Hancock, 2010). Such correlations may extend beyond the kinesin family into other systems, offering an explanation for domain conservation in families but with variance in linker lengths. In agreement with this hypothesis, metallothionein (MT) from plants and mammals has different linker lengths (longer in plants and shorter in mammalian), suggesting a possible explanation for the different metal-binding properties of the two-domain MT (Ngu et al., 2009).

### Difference in Sequence

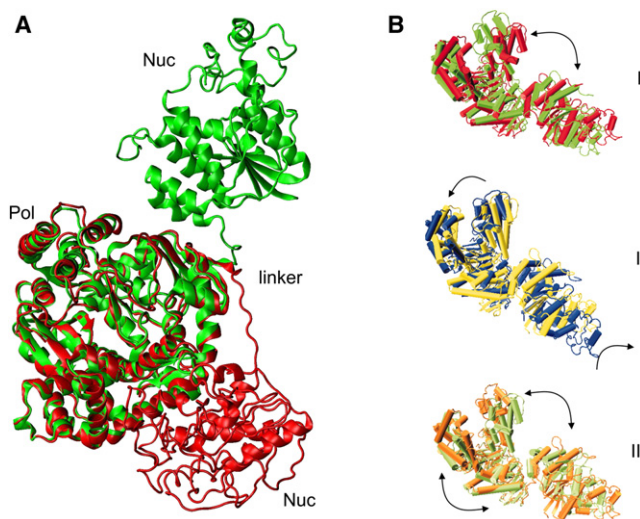
Sequences are important in the linker itself and in regions adjoining the linker. As comprehensively shown for the LacI/GalR bacterial family, even substitution of similar residues within the linkers may have dramatic functional consequences. Moreover, even substitutions at the same conserved positions may alter the function of a linker to different extents (Tungtur et al., 2011). An example for the role of sequences adjoining the linker is the neural cell adhesion molecule (NCAM), which is the primary substrate of the polysialyltransferases (polySTs). The first fibronectin type III repeat (FN1) of NCAM is required for polysialylation of the *N*-glycans on the adjacent immunoglobulin-like domain (Ig5), and acidic residues on the surface of FN1 were shown to play a role in polyST recognition (Thompson et al., 2011). Sequences adjoining linkers can play a role through the interactions they undergo with the linker during allosteric propagation.

The definition of linkers in the literature varies, and depending on size and conformations, linkers can also be domains that have roles other than transmitting signals, such as in the case of transcription factor IIB (TFIIB), the cyclin-dependent kinase 5 (Cdk5), or in the case of TGF- $\beta$ , where the linker presents a recognition motif. We propose that linkers should be classified according to their dynamic roles, rather than their lengths, or secondary structures. We further note that the mechanisms proposed here might not necessarily hold for all proteins, particularly for proteins that are rigid and consist of single domain chains.

### Conclusions

On its own, conformational flexibility of proteins does not constitute a practical solution for a cell: to present adequate biological responses to the changing environment, the cell needs: (i) preferred sampling of conformations that are relevant for function; and (ii) short timescales from the triggering event to the response, in particular to achieve high efficiencies of enzymatic metabolic reactions. For signaling and activation pathways, the difference between responses at a picosecond versus millisecond timescale to an activation event may not be significant. Because the majority of proteins consist of multiple domains, signals typically have to be communicated among domains, and this means that the information related to events (e.g., binding, covalent changes) in one domain has to be propagated quickly to other domains. Catalytic and effector sites are often separated by large distances, which allow larger complexity and tuning in regulation; however, this also emphasizes time consideration. Signals convert an “inactive” conformation to an “active” one. The timescales of conformational conversions are a function of the threshold between the different conformational states. The question arises how the protein overcomes the barriers separating the states to achieve fast response times. Here, we highlight two points: first, the linkers connecting domains are not only flexible but are allosterically regulated; and second, the sequence of the linkers and of the residues in contact between linkers and adjoining domains encodes successive states, through which the signals travel as illustrated by different conformational states encoded in Taq polymerase (Bu et al., 2005; Eom et al., 1996; Kim et al., 1995; Murali et al., 1998) (Figure 3). These sequences define pathways in which the consecutive barriers are lower, thereby making the communication (and, thus, the cooperative response) faster





**Figure 3. Crystal Structures and Neutron Spin-Echo Spectroscopy Revealed that the Different Conformational States Are Encoded in the Taq Polymerase**

Different conformational states that are encoded in the Taq polymerases have been revealed by crystal structures and neutron spin-echo spectroscopy. Taq polymerase has two large domains, the polymerase domain (Pol) and a structure-specific exonuclease (Nuc) domain, which are connected by a spring-like linker (Bu et al., 2005; Eom et al., 1996; Kim et al., 1995; Murali et al., 1998).

(A) In the active conformation (green ribbon) (Eom et al., 1996), the Pol and Nuc domains are separated. The inhibitory conformation was cocrystallized with an antibody (Murali et al., 1998), with the Nuc domain swinging back to bind with the Pol domain (red ribbon).

(B) Several encoded dynamic states (I, II, and III) have been observed by neutron spin-echo spectroscopy (Bu et al., 2005). The dynamic motion I corresponds to large conformation changes connecting the active and inhibitory conformations in (A).

The figures in (B) were taken from Bu et al. (2005) with permission (Copyright 2005, National Academy of Sciences, USA).

(Figure 2). Although there is experimental and computational evidence that validates the allosteric behavior of linkers, the concept that the sequences encode a series of states is more difficult to validate experimentally (Masterson et al., 2011). Pre-encoding a series of states with lower barriers between them facilitates preferred propagation pathways between favored states through which the population shift takes place. Coming back to the analogy of propagating waves following a perturbation in a body of water (del Sol et al., 2009), the waves will largely flow along those routes, in which single successive impediments are lower, rather than through fewer but higher obstacles. Thus, whereas all states preexist, the distributions of the populations change as the signals propagate: states that are lowly populated when the signal is at one point can become highly populated as the signal progresses. This suggests that evolution not only encodes states for direct function but also propagation pathways for cellular response. This is true not only for linkers; the arguments outlined here similarly hold for protein cores and particularly for flexible loops, which can mediate information transfer across the protein. This is further supported by studies illustrating coevolution of residues within a protein family, some of which are in linkers (Smock et al., 2010). Such groups of coevolving residues may pre-encode successive states.

## ACKNOWLEDGMENTS

T.H. acknowledges support from TUBA (Turkish Academy of Sciences) and State Planning organization DPT (2009K120520). This project has been funded in whole or in part with Federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported (in part) by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research.

## REFERENCES

- Apic, G., Huber, W., and Teichmann, S.A. (2003). Multi-domain protein families and domain pairs: comparison with known structures and a random model of domain recombination. *J. Struct. Funct. Genomics* 4, 67–78.
- Aravind, L., and Subramanian, G. (1999). Origin of multicellular eukaryotes—insights from proteome comparisons. *Curr. Opin. Genet. Dev.* 9, 688–694.
- Beckham, G.T., Bomble, Y.J., Matthews, J.F., Taylor, C.B., Resch, M.G., Yarbrough, J.M., Decker, S.R., Bu, L., Zhao, X., McCabe, C., et al. (2010). The O-glycosylated linker from the *Trichoderma reesei* Family 7 cellulase is a flexible, disordered protein. *Biophys. J.* 99, 3773–3781.
- Boehr, D.D. (2009). During transitions proteins make fleeting bonds. *Cell* 139, 1049–1051.
- Boehr, D.D., McElheny, D., Dyson, H.J., and Wright, P.E. (2006). The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* 313, 1638–1642.
- Boehr, D.D., Nussinov, R., and Wright, P.E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* 5, 789–796.
- Boehr, D.D., McElheny, D., Dyson, H.J., and Wright, P.E. (2010). Millisecond timescale fluctuations in dihydrofolate reductase are exquisitely sensitive to the bound ligands. *Proc. Natl. Acad. Sci. USA* 107, 1373–1378.
- Bruning, J.B., Parent, A.A., Gil, G., Zhao, M., Nowak, J., Pace, M.C., Smith, C.L., Afonine, P.V., Adams, P.D., Katzenellenbogen, J.A., and Nettles, K.W. (2010). Coupling of receptor conformation and ligand orientation determine graded activity. *Nat. Chem. Biol.* 6, 837–843.
- Brüschweiler, S., Schanda, P., Klotz, K., Brutscher, B., Kontaxis, G., Konrat, R., and Tollinger, M. (2009). Direct observation of the dynamic process underlying allosteric signal transmission. *J. Am. Chem. Soc.* 131, 3063–3068.
- Bu, Z., Biehl, R., Monkenbusch, M., Richter, D., and Callaway, D.J. (2005). Coupled protein domain motion in Taq polymerase revealed by neutron spin-echo spectroscopy. *Proc. Natl. Acad. Sci. USA* 102, 17646–17651.
- Cardozo, T., and Pagano, M. (2004). The SCF ubiquitin ligase: insights into a molecular machine. *Nat. Rev. Mol. Cell Biol.* 5, 739–751.
- Chang, J.T., Carvalho, C., Mori, S., Bild, A.H., Gatz, M.L., Wang, Q., Lucas, J.E., Potti, A., Febbo, P.G., West, M., and Nevins, J.R. (2009). A genomic strategy to elucidate modules of oncogenic pathway signaling networks. *Mol. Cell* 34, 104–114.
- Chen, J.X., and Kapral, R. (2011). Mesoscopic dynamics of diffusion-influenced enzyme kinetics. *J. Chem. Phys.* 134, 044503.
- Chen, W., and Roeder, R.G. (2007). The Mediator subunit MED1/TRAP220 is required for optimal glucocorticoid receptor-mediated transcription activation. *Nucleic Acids Res.* 35, 6161–6169.
- Chen, W., Rogatsky, I., and Garabedian, M.J. (2006). MED14 and MED1 differentially regulate target-specific gene activation by the glucocorticoid receptor. *Mol. Endocrinol.* 20, 560–572.
- Chong, P.A., Lin, H., Wrana, J.L., and Forman-Kay, J.D. (2010). Coupling of tandem Smad ubiquitination regulatory factor (Smurf) WW domains modulates target specificity. *Proc. Natl. Acad. Sci. USA* 107, 18404–18409.
- Cohen-Gihon, I., Sharan, R., and Nussinov, R. (2011). Processes of fungal proteome evolution and gain of function: gene duplication and domain rearrangement. *Phys. Biol.* 8, 035009.



- Csermely, P., Palotai, R., and Nussinov, R. (2010). Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem. Sci.* 35, 539–546.
- Cui, Q., and Karplus, M. (2008). Allostery and cooperativity revisited. *Protein Sci.* 17, 1295–1307.
- Cyrus, K., Wehenkel, M., Choi, E.Y., Han, H.J., Lee, H., Swanson, H., and Kim, K.B. (2011). Impact of linker length on the activity of PROTACs. *Mol. Biosyst.* 7, 359–364.
- del Sol, A., Tsai, C.J., Ma, B., and Nussinov, R. (2009). The origin of allosteric functional modulation: multiple pre-existing pathways. *Structure* 17, 1042–1050.
- Duda, D.M., Borg, L.A., Scott, D.C., Hunt, H.W., Hammel, M., and Schulman, B.A. (2008). Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* 134, 995–1006.
- Durante-Rodríguez, G., Valderrama, J.A., Mancheño, J.M., Rivas, G., Alfonso, C., Arias-Palomo, E., Llorca, O., García, J.L., Díaz, E., and Carmona, M. (2010). Biochemical characterization of the transcriptional regulator BzdR from *Azoarcus* sp. *CIB. J. Biol. Chem.* 285, 35694–35705.
- Echeverría, C., and Kapral, R. (2010). Macromolecular dynamics in crowded environments. *J. Chem. Phys.* 132, 104902.
- Eom, S.H., Wang, J., and Steitz, T.A. (1996). Structure of Taq polymerase with DNA at the polymerase active site. *Nature* 382, 278–281.
- Formanek, M.S., Ma, L., and Cui, Q. (2006). Reconciling the “old” and “new” views of protein allostery: a molecular simulation study of chemotaxis Y protein (CheY). *Proteins* 63, 846–867.
- Frauenfelder, H., Sligar, S.G., and Wolynes, P.G. (1991). The energy landscapes and motions of proteins. *Science* 254, 1598–1603.
- Friedmann, D.R., Wilson, J.J., and Kovall, R.A. (2008). RAM-induced allostery facilitates assembly of a notch pathway active transcription complex. *J. Biol. Chem.* 283, 14781–14791.
- Gardino, A.K., Villali, J., Kivenson, A., Lei, M., Liu, C.F., Steindel, P., Eisenmesser, E.Z., Labeikovsky, W., Wolf-Watz, M., Clarkson, M.W., and Kern, D. (2009). Transient non-native hydrogen bonds promote activation of a signaling protein. *Cell* 139, 1109–1118.
- Ge, K., Cho, Y.W., Guo, H., Hong, T.B., Guermah, M., Ito, M., Yu, H., Kalkum, M., and Roeder, R.G. (2008). Alternative mechanisms by which mediator subunit MED1/TRAP220 regulates peroxisome proliferator-activated receptor gamma-stimulated adipogenesis and target gene expression. *Mol. Cell. Biol.* 28, 1081–1091.
- George, R.A., and Heringa, J. (2002). An analysis of protein domain linkers: their classification and role in protein folding. *Protein Eng.* 15, 871–879.
- Gielen, M., Siegler Retchless, B., Mony, L., Johnson, J.W., and Paoletti, P. (2009). Mechanism of differential control of NMDA receptor activity by NR2 subunits. *Nature* 459, 703–707.
- Gokhale, R.S., and Khosla, C. (2000). Role of linkers in communication between protein modules. *Curr. Opin. Chem. Biol.* 4, 22–27.
- Goodey, N.M., and Benkovic, S.J. (2008). Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* 4, 474–482.
- Gronemeyer, H., and Bourguet, W. (2009). Allosteric effects govern nuclear receptor action: DNA appears as a player. *Sci. Signal.* 2, pe34.
- Gross-Langenhoff, M., Hofbauer, K., Weber, J., Schultz, A., and Schultz, J.E. (2006). cAMP is a ligand for the tandem GAF domain of human phosphodiesterase 10 and cGMP for the tandem GAF domain of phosphodiesterase 11. *J. Biol. Chem.* 281, 2841–2846.
- Guglielmi, B., van Berkum, N.L., Klapholz, B., Bijma, T., Boube, M., Boschiero, C., Bourbon, H.M., Holstege, F.C., and Werner, M. (2004). A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res.* 32, 5379–5391.
- Gunasekaran, K., Ma, B., and Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? *Proteins* 57, 433–443.
- Haliloglu, T., and Erman, B. (2009). Analysis of correlations between energy and residue fluctuations in native proteins and determination of specific sites for binding. *Phys. Rev. Lett.* 102, 088103.
- Hofbauer, K., Schultz, A., and Schultz, J.E. (2008). Functional chimeras of the phosphodiesterase 5 and 10 tandem GAF domains. *J. Biol. Chem.* 283, 25164–25170.
- Hollins, B., Kuravi, S., Digby, G.J., and Lambert, N.A. (2009). The c-terminus of GRK3 indicates rapid dissociation of G protein heterotrimers. *Cell. Signal.* 21, 1015–1021.
- Huang, Z., Zhu, L., Cao, Y., Wu, G., Liu, X., Chen, Y., Wang, Q., Shi, T., Zhao, Y., Wang, Y., et al. (2011). ASD: a comprehensive database of allosteric proteins and modulators. *Nucleic Acids Res.* 39 (Database issue), D663–D669.
- Ito, M., Yuan, C.X., Malik, S., Gu, W., Fondell, J.D., Yamamura, S., Fu, Z.Y., Zhang, X., Qin, J., and Roeder, R.G. (1999). Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol. Cell* 3, 361–370.
- Itoh, K., and Sasai, M. (2010). Entropic mechanism of large fluctuation in allosteric transition. *Proc. Natl. Acad. Sci. USA* 107, 7775–7780.
- Joseph, R.E., Xie, Q., and Andreotti, A.H. (2010). Identification of an allosteric signaling network within Tec family kinases. *J. Mol. Biol.* 403, 231–242.
- Kalodimos, C.G. (2011). NMR reveals novel mechanisms of protein activity regulation. *Protein Sci.* 20, 773–782.
- Kang, Y.K., Guermah, M., Yuan, C.X., and Roeder, R.G. (2002). The TRAP/Mediator coactivator complex interacts directly with estrogen receptors alpha and beta through the TRAP220 subunit and directly enhances estrogen receptor function in vitro. *Proc. Natl. Acad. Sci. USA* 99, 2642–2647.
- Kannan, N., Wu, J., Anand, G.S., Yooseph, S., Neuwald, A.F., Venter, J.C., and Taylor, S.S. (2007). Evolution of allostery in the cyclic nucleotide binding module. *Genome Biol.* 8, R264.
- Kar, G., Keskin, O., Gursoy, A., and Nussinov, R. (2010). Allostery and population shift in drug discovery. *Curr. Opin. Pharmacol.* 10, 715–722.
- Kenakin, T. (2010). G protein coupled receptors as allosteric proteins and the role of allosteric modulators. *J. Recept. Signal Transduct. Res.* 30, 313–321.
- Kenakin, T., and Miller, L.J. (2010). Seven transmembrane receptors as shape-shifting proteins: the impact of allosteric modulation and functional selectivity on new drug discovery. *Pharmacol. Rev.* 62, 265–304.
- Kern, D., and Zuiderweg, E.R. (2003). The role of dynamics in allosteric regulation. *Curr. Opin. Struct. Biol.* 13, 748–757.
- Keskin, O. (2007). Binding induced conformational changes of proteins correlate with their intrinsic fluctuations: a case study of antibodies. *BMC Struct. Biol.* 7, 31.
- Kim, Y., Eom, S.H., Wang, J., Lee, D.S., Suh, S.W., and Steitz, T.A. (1995). Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* 376, 612–616.
- Lee, H.K., Park, U.H., Kim, E.J., and Um, S.J. (2007). MED25 is distinct from TRAP220/MED1 in cooperating with CBP for retinoid receptor activation. *EMBO J.* 26, 3545–3557.
- Li, H., Gade, P., Nallar, S.C., Raha, A., Roy, S.K., Karra, S., Reddy, J.K., Reddy, S.P., and Kalvakolanu, D.V. (2008). The Med1 subunit of transcriptional mediator plays a central role in regulating CCAAT/enhancer-binding protein-beta-driven transcription in response to interferon-gamma. *J. Biol. Chem.* 283, 13077–13086.
- Liu, J., and Nussinov, R. (2009). The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: conformational control of substrate orientation. *PLoS Comput. Biol.* 5, e1000527.
- Liu, J., and Nussinov, R. (2010a). Molecular dynamics reveal the essential role of linker motions in the function of cullin-RING E3 ligases. *J. Mol. Biol.* 396, 1508–1523.
- Liu, J., and Nussinov, R. (2010b). Rbx1 flexible linker facilitates cullin-RING ligase function before neddylation and after deneddylation. *Biophys. J.* 99, 736–744.

- Ma, B., and Nussinov, R. (2009). Amplification of signaling via cellular allosteric relay and protein disorder. *Proc. Natl. Acad. Sci. USA* 106, 6887–6888.
- Ma, B., and Nussinov, R. (2010). Enzyme dynamics point to stepwise conformational selection in catalysis. *Curr. Opin. Chem. Biol.* 14, 652–659.
- Ma, B., Kumar, S., Tsai, C.J., and Nussinov, R. (1999). Folding funnels and binding mechanisms. *Protein Eng.* 12, 713–720.
- Malik, S., Wallberg, A.E., Kang, Y.K., and Roeder, R.G. (2002). TRAP/SMCC/mediator-dependent transcriptional activation from DNA and chromatin templates by orphan nuclear receptor hepatocyte nuclear factor 4. *Mol. Cell. Biol.* 22, 5626–5637.
- Malik, S., Guermah, M., Yuan, C.X., Wu, W.Z., Yamamura, S., and Roeder, R.G. (2004). Structural and functional organization of TRAP220, the TRAP/mediator subunit that is targeted by nuclear receptors. *Mol. Cell. Biol.* 24, 8244–8254.
- Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., and Eisenberg, D. (1999). Detecting protein function and protein-protein interactions from genome sequences. *Science* 285, 751–753.
- Masterson, L.R., Shi, L., Metcalfe, E., Gao, J., Taylor, S.S., and Veglia, G. (2011). Dynamically committed, uncommitted, and quenched states encoded in protein kinase A revealed by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* 108, 6969–6974.
- Meijssing, S.H., Pufall, M.A., So, A.Y., Bates, D.L., Chen, L., and Yamamoto, K.R. (2009). DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* 324, 407–410.
- Meyer, K.D., Lin, S.C., Bernecky, C., Gao, Y., and Taatjes, D.J. (2010). p53 activates transcription by directing structural shifts in Mediator. *Nat. Struct. Mol. Biol.* 17, 753–760.
- Mittermaier, A.K., and Kay, L.E. (2009). Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.* 34, 601–611.
- Murali, R., Sharkey, D.J., Daiss, J.L., and Murthy, H.M. (1998). Crystal structure of Taq DNA polymerase in complex with an inhibitory Fab: the Fab is directed against an intermediate in the helix-coil dynamics of the enzyme. *Proc. Natl. Acad. Sci. USA* 95, 12562–12567.
- Natarajan, K., Jackson, B.M., Zhou, H., Winston, F., and Hinnebusch, A.G. (1999). Transcriptional activation by Gcn4p involves independent interactions with the SWI/SNF complex and the SRB/mediator. *Mol. Cell* 4, 657–664.
- Ngu, T.T., Lee, J.A., Rushton, M.K., and Stillman, M.J. (2009). Arsenic metalation of seaweed *Fucus vesiculosus* metallothionein: the importance of the interdomain linker in metallothionein. *Biochemistry* 48, 8806–8816.
- Okazaki, K., and Takada, S. (2008). Dynamic energy landscape view of coupled binding and protein conformational change: induced-fit versus population-shift mechanisms. *Proc. Natl. Acad. Sci. USA* 105, 11182–11187.
- Pandit, J., Forman, M.D., Fennell, K.F., Dillman, K.S., and Menniti, F.S. (2009). Mechanism for the allosteric regulation of phosphodiesterase 2A deduced from the X-ray structure of a near full-length construct. *Proc. Natl. Acad. Sci. USA* 106, 18225–18230.
- Paris, L.L., Hu, J., Galan, J., Ong, S.S., Martin, V.A., Ma, H., Tao, W.A., Harrison, M.L., and Geahlen, R.L. (2010). Regulation of Syk by phosphorylation on serine in the linker insert. *J. Biol. Chem.* 285, 39844–39854.
- Pawson, T. (2003). Organization of cell-regulatory systems through modular-protein-interaction domains. *Philos. Transact. A Math. Phys. Eng. Sci.* 361, 1251–1262.
- Peisajovich, S.G., Garbarino, J.E., Wei, P., and Lim, W.A. (2010). Rapid diversification of cell signaling phenotypes by modular domain recombination. *Science* 328, 368–372.
- Petit, C.M., Zhang, J., Sapienza, P.J., Fuentes, E.J., and Lee, A.L. (2009). Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA* 106, 18249–18254.
- Piazza, F., and Sanejouand, Y.H. (2008). Discrete breathers in protein structures. *Phys. Biol.* 5, 026001.
- Piazza, F., and Sanejouand, Y.H. (2009). Long-range energy transfer in proteins. *Phys. Biol.* 6, 046014.
- Popovych, N., Tzeng, S.R., Tonelli, M., Ebright, R.H., and Kalodimos, C.G. (2009). Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. *Proc. Natl. Acad. Sci. USA* 106, 6927–6932.
- Rabone, J., Yue, Y.F., Chong, S.Y., Stylianou, K.C., Bacsá, J., Bradshaw, D., Darling, G.R., Berry, N.G., Khimyak, Y.Z., Ganin, A.Y., et al. (2010). An adaptable peptide-based porous material. *Science* 329, 1053–1057.
- Ryan, P.E., Sivadasan-Nair, N., Nau, M.M., Nicholas, S., and Lipkowitz, S. (2010). The N terminus of Cbl-c regulates ubiquitin ligase activity by modulating affinity for the ubiquitin-conjugating enzyme. *J. Biol. Chem.* 285, 23687–23698.
- Schulman, B.A., Carrano, A.C., Jeffrey, P.D., Bowen, Z., Kinnucan, E.R., Finnin, M.S., Elledge, S.J., Harper, J.W., Pagano, M., and Pavletich, N.P. (2000). Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex. *Nature* 408, 381–386.
- Shastri, S., and Hancock, W.O. (2010). Neck linker length determines the degree of processivity in kinesin-1 and kinesin-2 motors. *Curr. Biol.* 20, 939–943.
- Smock, R.G., Rivoire, O., Russ, W.P., Swain, J.F., Leibler, S., Ranganathan, R., and Gierasch, L.M. (2010). An interdomain sector mediating allostery in Hsp70 molecular chaperones. *Mol. Syst. Biol.* 6, 414.
- Strickland, D., Moffat, K., and Sosnick, T.R. (2008). Light-activated DNA binding in a designed allosteric protein. *Proc. Natl. Acad. Sci. USA* 105, 10709–10714.
- Stumpf, M., Waskow, C., Krötschel, M., van Essen, D., Rodriguez, P., Zhang, X., Guyot, B., Roeder, R.G., and Borggrefe, T. (2006). The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proc. Natl. Acad. Sci. USA* 103, 18504–18509.
- Swain, J.F., and Gierasch, L.M. (2006). The changing landscape of protein allostery. *Curr. Opin. Struct. Biol.* 16, 102–108.
- Swanson, M.J., Qiu, H., Sumibcay, L., Krueger, A., Kim, S.J., Natarajan, K., Yoon, S., and Hinnebusch, A.G. (2003). A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Mol. Cell. Biol.* 23, 2800–2820.
- Taraska, J.W., and Zagotta, W.N. (2007). Structural dynamics in the gating ring of cyclic nucleotide-gated ion channels. *Nat. Struct. Mol. Biol.* 14, 854–860.
- Thompson, M.G., Foley, D.A., Swartzentruber, K.G., and Colley, K.J. (2011). Sequences at the interface of the fifth immunoglobulin domain and first fibronectin type III repeat of the neural cell adhesion molecule are critical for its polysialylation. *J. Biol. Chem.* 286, 4525–4534.
- Tobi, D., and Bahar, I. (2005). Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. *Proc. Natl. Acad. Sci. USA* 102, 18908–18913.
- Tsai, C.J., Ma, B., and Nussinov, R. (1999a). Folding and binding cascades: shifts in energy landscapes. *Proc. Natl. Acad. Sci. USA* 96, 9970–9972.
- Tsai, C.J., Kumar, S., Ma, B., and Nussinov, R. (1999b). Folding funnels, binding funnels, and protein function. *Protein Sci.* 8, 1181–1190.
- Tsai, C.J., del Sol, A., and Nussinov, R. (2008). Allostery: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol.* 378, 1–11.
- Tsai, C.J., del Sol, A., and Nussinov, R. (2009). Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms. *Mol. Biosyst.* 5, 207–216.
- Tsai, C.J., Ma, B., Sham, Y.Y., Kumar, S., and Nussinov, R. (2001). Structured disorder and conformational selection. *Proteins* 44, 418–427.
- Tungtur, S., Parente, D.J., and Swint-Kruse, L. (2011). Functionally important positions can comprise the majority of a protein's architecture. *Proteins* 79, 1589–1608.
- Tzeng, S.R., and Kalodimos, C.G. (2009). Dynamic activation of an allosteric regulatory protein. *Nature* 462, 368–372.
- Udayakumar, T.S., Belakavadi, M., Choi, K.H., Pandey, P.K., and Fondelli, J.D. (2006). Regulation of Aurora-A kinase gene expression via GABP recruitment of TRAP220/MED1. *J. Biol. Chem.* 281, 14691–14699.

- Villali, J., and Kern, D. (2010). Choreographing an enzyme's dance. *Curr. Opin. Chem. Biol.* 14, 636–643.
- Volkman, B.F., Lipson, D., Wemmer, D.E., and Kern, D. (2001). Two-state allosteric behavior in a single-domain signaling protein. *Science* 291, 2429–2433.
- Wallberg, A.E., Yamamura, S., Malik, S., Spiegelman, B.M., and Roeder, R.G. (2003). Coordination of p300-mediated chromatin remodeling and TRAP/mediator function through coactivator PGC-1 $\alpha$ . *Mol. Cell* 12, 1137–1149.
- Wang, F., Mei, Z., Qi, Y., Yan, C., Hu, Q., Wang, J., and Shi, Y. (2011). Structure and mechanism of the hexameric MecA-ClpC molecular machine. *Nature* 471, 331–335.
- Whitford, P.C., Onuchic, J.N., and Wolynes, P.G. (2008). Energy landscape along an enzymatic reaction trajectory: hinges or cracks? *HFSP J.* 2, 61–64.
- Wickström, S.A., and Fässler, R. (2011). Regulation of membrane traffic by integrin signaling. *Trends Cell Biol.* 21, 266–273.
- Wriggers, W., Chakravarty, S., and Jennings, P.A. (2005). Control of protein functional dynamics by peptide linkers. *Biopolymers* 80, 736–746.
- Yehekel, A., Haliloglu, T., and Ben-Tal, N. (2010). Independent and cooperative motions of the Kv1.2 channel: voltage sensing and gating. *Biophys. J.* 98, 2179–2188.
- Zhan, H., Taraban, M., Trehella, J., and Swint-Kruse, L. (2008). Subdividing repressor function: DNA binding affinity, selectivity, and allostery can be altered by amino acid substitution of nonconserved residues in a LacI/GalR homologue. *Biochemistry* 47, 8058–8069.
- Zhang, F., Sumibcay, L., Hinnebusch, A.G., and Swanson, M.J. (2004). A triad of subunits from the Gal11/tail domain of Srb mediator is an in vivo target of transcriptional activator Gcn4p. *Mol. Cell. Biol.* 24, 6871–6886.
- Zhang, J., Sapienza, P.J., Ke, H., Chang, A., Hengel, S.R., Wang, H., Phillips, G.N., and Lee, A.L. (2010). Crystallographic and nuclear magnetic resonance evaluation of the impact of peptide binding to the second PDZ domain of protein tyrosine phosphatase 1E. *Biochemistry* 49, 9280–9291.
- Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., et al. (2002). Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* 416, 703–709.
- Zhu, J., Recio-Pinto, E., Hartwig, T., Sellers, W., Yan, J., and Thornhill, W.B. (2009). The Kv1.2 potassium channel: the position of an N-glycan on the extracellular linkers affects its protein expression and function. *Brain Res.* 1251, 16–29.
- Zhu, L., Chen, L., Zhou, X.M., Zhang, Y.Y., Zhang, Y.J., Zhao, J., Ji, S.R., Wu, J.W., and Wu, Y. (2011). Structural insights into the architecture and allostery of full-length AMP-activated protein kinase. *Structure* 19, 515–522.
- Zhuravlev, P.I., and Papoian, G.A. (2010). Protein functional landscapes, dynamics, allostery: a tortuous path towards a universal theoretical framework. *Q. Rev. Biophys.* 43, 295–332.
- Zhuravleva, A., and Gierasch, L.M. (2011). Allosteric signal transmission in the nucleotide-binding domain of 70-kDa heat shock protein (Hsp70) molecular chaperones. *Proc. Natl. Acad. Sci. USA* 108, 6987–6992.